A FUZZY RULE BASED INFERENCE SYSTEM FOR EARLY DEBT COLLECTION

Sezi Cevik ONAR, Basar OZTAYSI, Cengiz KAHRAMAN[*]

*Management Faculty, Industrial Engineering Department,
Istanbul Technical University, 34367, Macka, Istanbul, Turkey*

Abstract. Nowadays, unpaid invoices and unpaid credits are becoming more and more common. Large amounts of data regarding these debts are collected and stored by debt collection agencies. Early debt collection processes aim at collecting payments from creditors or debtors before the legal procedure starts. In order to be successful and be able to collect maximum debts, collection agencies need to use their human resources efficiently and communicate with the customers via the most convenient channel that leads to minimum costs. However, achieving these goals need processing, analyzing and evaluating customer data and inferring the right actions instantaneously. In this study, fuzzy inference based intelligent systems are used to empower early debt collection processes using the principles of data science. In the paper, an early debt collection system composed of three different Fuzzy Inference Systems (FIS), one for credit debts, one for credit card debts, and one for invoices, is developed. These systems use different inputs such as *amount of loan*, *wealth of debtor*, *part history of debtor*, *amount of other debts*, *active customer since*, *credit limit*, and *criticality* to determine the output *possibility of repaying the debt*. This output is later used to determine the most convenient communication channel and communication activity profile.

Keywords: fuzzy inference system, early debt collection, credit, credit card, overdraft, invoice.

JEL Classification: D81, H63.

Introduction

Debt Collection Systems are very important for the survival of the organizations. Unfortunately, in the current global recessionary climate, unpaid invoices and unpaid credits are becoming more and more common. Companies may have sold goods/services and invoiced them to their customers but left waiting for the payment, or banks may have given credits but couldn't get the imbursements. These kinds of situations can seriously affect cash flows, turnover, credit ratings and even business' reputation. Debt collection processes take in case of unpaid invoices or credit imbursements to gather the unpaid amount. Debt collection

---

*Corresponding author. E-mail: *kahramanc@itu.edu.tr*

agencies have large amount of data sets from banks, GSM operators, and electricity-natural gas-internet service providers. Hence, a successful debt collection system necessitates the techniques of data science to be utilized. Data science refers to the process of learning and generalizable extraction of knowledge from data to gain useful predictions and insights. The data science technique used in this paper is Mamdani-type Fuzzy Inference System.

During the recent years, Turkey's credit applications have increased whereas interest rates are decreased; also competition among banks has increased which caused lower profitability. According to Turkish Banking Regulation and Supervision agency as of March 2013, the total amount of issued credits in Turkey is around 420 billion dollars and the total amount of past due payments are around 13 billion dollars. Collecting payments from the customers with a minimum cost before the legal procedure is crucial for the organizations since legal procedure is not only costly but also damages the customer relations. On the other hand, due to the inconvenient debt collection systems, there are debtors who directly face with legal processes without any warning from collection agencies and they have to compensate high interest rates. As a result, banks try to collect the debts before legal process starts and for this purpose they send e-mails, text messages, or call the debtors. Collection in litigation starts after 90 days past due. The collection systems trying to collect debts within 0–90 days are called early collection systems. In Turkey, there are currently some approaches that support early debt collection systems (0–90 days) but they are generally primitive and their efficiencies are low.

Either the organizations contact the customers on regular time bases with the same channels, or the call center agents use their own judgments for selecting any channel. Contacting on a regular base increases the cost of debt collection. On the other hand, inconsistencies occur when the call center agents use their own judgments and this damages customer relationships. Consequently, organizations need effective debt collection systems. On the other hand, in the literature only a few studies focused on early collection systems thus an intelligent early collection system will be beneficial for both organizations and academicians. The early debt collection system evaluates different aspects of the debt in order to define actions towards customers. In this process, the collection agents define the customers' possibility of repaying the debt with various criteria. Afterwards different actions are taken based on this possibility. This process contains high degree of subjectivity, vagueness and imprecision, since definition of high loan can be different for each collection agent. The fuzzy set theory enables us to deal with the problems of imprecisely defined objects (Ross 1995; Kahraman *et al.* 2010; Kaya *et al.* 2012; Cevik Onar *et al.* 2015; Cevik Onar, Ates 2008). Fuzzy inference systems formulate the mapping from the given inputs to an output via fuzzy logic. The decisions can be made based on the results of mapping processes. Fuzzy inference systems use fuzzy IF-THEN rules. Due to their concise form, fuzzy IF-THEN rules are often employed to capture the imprecise modes of reasoning that play an essential role in the human ability to make decisions in uncertain and imprecise environment (Tavana *et al.* 2013). Designing early debt collection system based on fuzzy rule based inference systems will limit the subjectivity, and increase the effectiveness.

This study proposes an effective fuzzy rule based early collection system by considering the complexity and the vagueness of the system. Although there are many fuzzy rule based

applications in the literature, none of them uses this method in early debt collection systems. Moreover, in order to improve effectiveness of the collection system, three different debt collection sub-systems are designed for different debt types.

To the best of our knowledge, FIS has not yet been used for financial purposes as it is in this paper. Humans have difficulty in making judgments when the number of inputs is more than two. In our study the number of inputs varies from four to five. Besides, each input is composed of at least five linguistic terms. This causes inconsistent judgments to be made by humans in inferential systems. Having many judgment makers in a debt collection center also causes inconsistent applications since each judgment maker may constitute his/her own independent judgment. This results in an increase in costs and customer complaints. Innovation of our study is to solve this complex and unhandled problem by a FIS approach. An in-depth system analysis has been conducted through many interviews and software has been developed, which suggests a certain communication channel and/or automatically communicates with debtors.

The rest of the paper is organized as follows. Section 1 gives a review of debt collection systems. Section 2 first gives a brief literature review of fuzzy rule based inference systems and then introduces the methodology. In Section 3, a real case study is given and the early debt collection systems are designed for three different debt types. Finally, the last section gives the conclusions and future research suggestions.

1. Debt collection systems

Companies and banks may try to handle debt collection processes by themselves or may outsource the work to a debt collection agency. A recent market research report on debt collection agencies in USA indicates that 9,599 agencies exist in USA market making a 13 billion dollars revenue (IBISworld 2014). Howard (2012) highlights the major problems of debt collection as follows (i) it takes more people/resources to collect (ii) debt is costing companies more than ever before (iii) the costs are escalating because debt collection is taking longer to retrieve. The results indicate that debt collection is very important for companies and should be handled very carefully.

Because of the legislations about debt collection such as fair debt collection practices act (FDSPA 2014) in USA, and policies designed by the Member States of European Union (Huls 1992) the collectors should be very attentive about their behaviors to the customers. A wrong communication style can also damage the relationship between companies and customers. Lund (2010) identifies the term Soft Debt Collection (SDC) as a process of understanding your customer's business and their reasons for not paying, and choosing the appropriate course of action. From the debt collectors' perspective, if the debtor is in trouble, it is vital not to be the last one in the queue, but, pushing too hard when they might only have short-term cash-flow problems could lose a customer in whom a lot of time and money is invested for acquisition and development. Pushing too hard can also tip a customer over the edge into insolvency, in which case you may receive nothing, whereas understanding their problems and getting proper actions may put the collector in a much better position.

With the emergence of information age, the collectors are using many information sources to get information about debtors and to reach them, such as internet, mobile phones, telephones, emails, voice messages (Hector 2011). Each of these technologies has different levels of automation. For instance, emails and text messages can be sent automatically and thus each of the channels causes different amount of costs. For example, an email message can easily be sent to a debtor automatically by a software program causing nearly zero costs, similarly an automatic text message can be sent to a debtor but this time for each message a fee should be paid to GSM operator. Most dramatically, to establish a phone call, the collector needs to pay a fee for the telephone service provider and also an operator is needed to handle the call. The selection of the most proper communication type directly affects the costs of the collector. Besides, which type of communication is preferred by the debtor, which type of message should be used for written communications, which tone of speaking should be used in phone calls, and at what time of the day should the debtor be called also affect the relationship between the debtor and the customer.

In the literature there are a number of studies which use analytical approaches to empower debt collection systems. In one of the most recent studies, Takahashi and Tsuda (2013) focus on the characteristics of bad customers in mail order industry. The authors build a system for identifying potential debtors using a random forest approach. The results of the study show that the location and monetary value of the transaction are two important factors for identifying potential debtors. Wang *et al.* (2013) focus on telecommunication market in Taiwan and propose a model to prevent companies from increasing bad debt and experiencing customer churn. The authors build a behavioral customer scoring model using a decision tree. The results of the practical application show that the total collection costs are reduced by 0.4% of annual revenue and significant reduction in customer churn caused by chaotic dunning strategies are realized.

Chen and Huang (2011) present a data mining approach to predict credit card customer behaviors. The authors use artificial neural networks and decision trees to predict regular patterns of consumption, payment/default, and bad debt. Abe *et al.* (2010) focus on optimally managing debt collection processes in financial institutions. The authors integrate data modeling and optimization techniques for debt collection optimization. The model is deployed to New York State Department of Taxation and Finance Office in December 2009. The monetary benefits of the systems are expected to be 100 million dollars in the first three years. Fei (2010) studies repayment behaviors of credit card holders. In their two-stage behavioral scoring model, they first use Chi-square automatic interaction detector (CHAID) and artificial neural networks for initial classification. Based on the results of the first stage, the important factors of classification are determined. Later these factors are used to build Data Envelopment Analysis (DEA) in the second stage. The aim of using DEA is to evaluate individual behavior scores in order to help banks reduce the cost of potential misclassification of customers.

Georgopoulos and Giannaropoulos (2007) focus on resource optimization of contact centers that interact with debtors for debt collection. The authors use Artificial Neural Networks (ANN) in order to develop a collection scoring system for debtors. Using these scores, the contact centers can better utilize their agents and optimize total collection costs. In another

study, Vecchio *et al.* (2006) investigate locating a bad debtor which attempts to escape a debt. In such cases, debt collection agencies search an address database for possible addresses of the debtor. The aim of the study is to determine the accurate address when there are multiple addresses returned from the database. The authors develop a MS Excel Macro based on Levenshtein Distance and Hierarchical Clustering and the accuracy of the algorithm is reported as 70.46%. Chin and Kotak (2006) present a case study about using rule-based decision engines to improve debt collection processes. The rule based system has been applied in a bank in nearly two years and the project team created more than 20.000 rules for the rule engine.

2. Rule based systems

Fuzzy rule based systems are composed of fuzzy rules with linguistic inputs and outputs to obtain a decision based on all the rules. There are two types of fuzzy inference systems (FISs): Sugeno and Mamdani inference systems. When the inputs are given, there are six steps to compute the output of the Mamdani FIS (1995): (1) determine a set of fuzzy rules, (2) fuzzify the inputs using the input membership functions, (3) combine the fuzzified inputs according to the fuzzy rules to establish a rule strength, (4) find the consequence of the rule by combining the rule strength and the output membership function, (5) combine the consequences to get an output distribution, and (6) defuzzify the output distribution.

The Sugeno FIS is quite similar to the Mamdani FIS. The primary difference is that the output consequence is not computed by clipping an output membership function at the rule strength. In fact, in the Sugeno FIS there is no output membership function at all. Instead the output is a crisp number computed by multiplying each input by a constant and then adding up the results. The Sugeno output membership functions may be either linear or constant.

A fuzzy rule base consists of a set of fuzzy IF-THEN rules. It is the core of the FIS in the sense that all other components, e.g. membership functions, have been employed to implement these rules in a reasonable, realistic and efficient manner. These fuzzy IF-THEN rules are utilized by the FIS to determine a mapping from fuzzy sets in the input universe of discourse $U \subset R^n$ to fuzzy sets in the output universe of discourse $V \subset R$, based on fuzzy logic principles. The fuzzy IF-THEN rules can be given as in Eq. (1) (Tavana *et al.* 2013):

$$R^{(l)} : IF\ x_1\ is\ F_1^l\ and\ \dots\ x_n\ is\ F_n^l,\ then\ y\ is\ G^l, \tag{1}$$

where F_Y^l and G^l are fuzzy sets in $U_i \subset R$, respectively and $x = (x_1, x_2, \dots, x_n)^T \in U$ and $y \in V$ are input and output linguistic variables of the FIS which belong to the input and output universes, respectively. l represents the number of rules in the fuzzy rule base, $l = 1, 2, \dots, M$.

To evaluate the disjunction of the rule inputs, an OR fuzzy operation is used. Typically, using the fuzzy operation union:

$$\mu_{A \cup B}(X) = \max\left[\mu_A(X), \mu_B(X)\right]. \tag{2}$$

Similarly, in order to evaluate the conjunction of the rule inputs, an AND fuzzy operation is applied to the intersection:

$$\mu_{A \cup B}(X) = \min\left[\mu_A(X), \mu_B(X)\right]. \tag{3}$$

The defuzzification of the union of the outputs can be made by using the centroid of area method which is given by Eq. (4).

$$COA = \frac{\int \mu_z(x) x\, dx}{\int \mu_z(x)\, dx}, \tag{4}$$

where μ_z is the membership function of the union of the outputs.

In Sugeno type FIS, the defuzzification is made by Eq. (5):

$$\text{Weighted Average} = \frac{\sum_{i=1}^{N} w_i z_i}{\sum_{i=1}^{N} w_i}, \tag{5}$$

where w_i is firing strength of rule i and z_i is the output of rule i, which can be calculated by $z_i = a_i x + b_i y + c_i$ for a two inputs rule. In this equation, a_i, b_i, and c_i are the user-defined constants.

Our work uses a Mamdani type FIS since it has widespread acceptance and well suited to human inputs. For more information on FIS, readers can read (Tay, Lim 2008a). FIS have been used in many application areas such as fuzzy assessment models for failure mode and effect analysis (FMEA) (Tay, Lim 2008b; Jee *et al.* 2015), prediction of stock market return and economic crisis based on adaptive network-based fuzzy inference system (ANFIS) (Boyacioglu, Avci 2010; Giovanis 2012), spontaneous volunteer management for disaster response (Oztaysi *et al.* 2013), supply chain risk management (Behret *et al.* 2011), performance prediction of solar thermal energy system (Yaici, Entchev 2016), and sustainability assessment of manufacturing processes (Rezvan *et al.* 2014).

A Mamdani type FIS with two rules is illustrated for the early debt collection problem in Figure 1. The inputs of the inference system for credit debts are amount of loan, wealth of the debtor, past history of the debtor, amount of other debts, and active customer since (how along he/she is our active customer) while the output is the possibility of repaying debt. The output is represented by the linguistic terms very low, low, low to medium, medium, medium to high, high and very high. As seen in Figure 1, the minimum membership degree among all inputs determines the output fuzzy set for each rule. Later, the union of all output fuzzy sets yields the output distribution. The defuzzification of the output distribution gives the possibility of repaying debt for those input values. The first line of Figure 1 presents a fuzzy rule composed of five inputs and one output while its second line represents another fuzzy rule. In the first rule, amount of loan is low, wealth of the debtor is high, past history of the debtor is low, amount of other debts is low, active customer since is high and then the possibility of repaying debt is high. Similarly, second rule yields the output possibility of repaying debt as very high. A certain case whose input values are known determines the membership degree for each input as shown Figure 1. The amount of each input gives the membership degree indicated by a dashed line. The shaded area for the output of each rule is determined by the minimum of input membership degrees. Later, the union of shaded areas gives the aggregated output that will be defuzzified by a defuzzification method to find the crisp output.
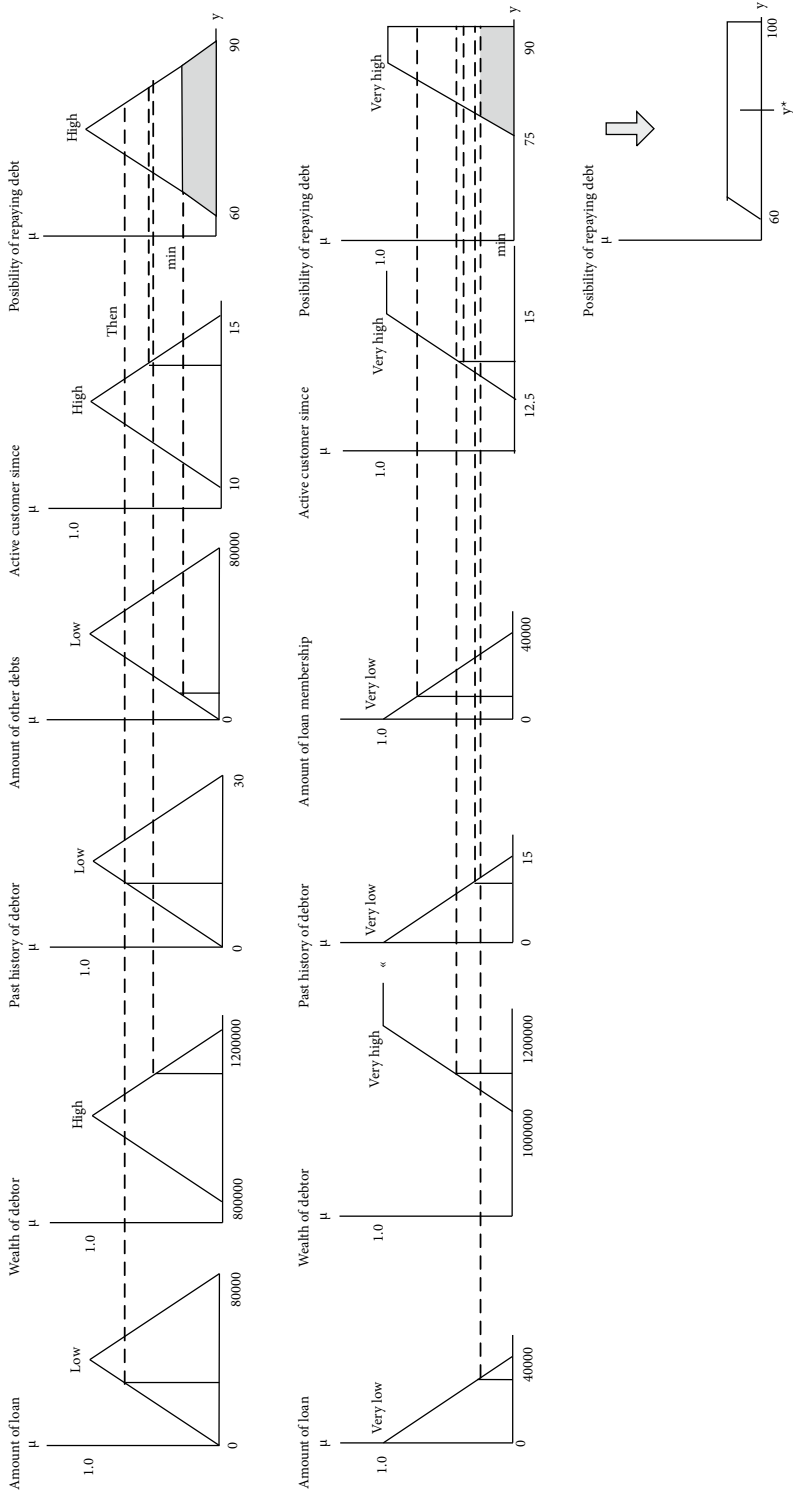
Figure 1. Fuzzy inference method

3. A real case study

A debt collection agency in Turkey wants to develop an intelligent early collection system. This firm is an expert agency on problematic debts and is one of Turkey's leading debt collection agencies. 9 out of 13 major banks, all GSM operators, and electricity providers are the customers of this firm.

In the present case, either the debt collection agency contacts with debtors in regular time intervals using the certain channels such as mobile phones, text messages, e-mail messages, and formal letters or the call center agents on behalf of the debt collection agency use their own instant judgments in selecting any of these channels. The selection of a communication channel without a detailed analysis of a certain case significantly increases the cost of the whole process. For instance, using mobile phone for contacting a debtor whose possibility of paying the debt is very low, whereas an e-mail message would be sufficient, will clearly cause an unnecessary increase in communication cost. On the contrary, contacting a debtor with an e-mail message whose possibility of paying the debt is very high, whereas a mobile phone would be necessary, may cause the debtor not to pay his/her debt. Besides, when the huge number of debtors is considered, the total cost of communication may reach to million dollars.

The fuzzy inference system has been developed based on expert opinions. A team of eight experts (five of the experts have field experience; three of them are academicians) had ten meetings. The time spent in each meeting was between 1.5–3 hours.

Based on data analysis and expert opinions, three different debt types have been identified. These debt types are credit debts, credit card and overdraft debts, and invoices. Each of these debt types necessitates different FIS based debt collection systems to be generated. In this system totally 33.320 rules have been generated, that is the sum of 15.680 rules for credit debts, 15.680 rules for credit card and overdraft debts, and 1.960 rules for invoices. 15.680 rules for credit debts is the result of the multiplication $7 \times 7 \times 8 \times 5 \times 8$ where *amount of loan is represented by 7* linguistic terms; *wealth of debtor* is represented by 8 linguistic terms including "no information"; *amount of other debts* is represented by 7 linguistic terms; *past history of debtor* is represented by 5 linguistic terms; and *active customer since is represented* by 8 linguistic terms including *"no information"*. 15,680 rules for credit card is obtained by multiplication $7 \times 8 \times 8 \times 5 \times 7$ where *amount of loan* is represented by 7 linguistic terms; *active customer since* is represented by 8 linguistic terms including *"no information"; amount of other debts* is represented by 8 linguistic terms including *"no information"; past history of debtor* is represented by 5 linguistic terms; and *credit limit* is represented by 7 linguistic terms. 1,960 is the result of the multiplication $7 \times 8 \times 5 \times 7$ where *amount of loan* 7 linguistic terms; active *customer since* 8 linguistic terms including *"no information"; past history of debtor* has 7 linguistic terms; and *criticality* has 5 linguistic terms.

The whole picture of the application is given in Figure 2. It shows the three categories of the debts, which are obtained from the debt collection database. When a new debt is introduced to the debt collection database, first its category is defined and based on this category the inference system that will handle this debt is determined. The information regarding the debt and debtor is obtained from the debt collection database and entered to the relevant
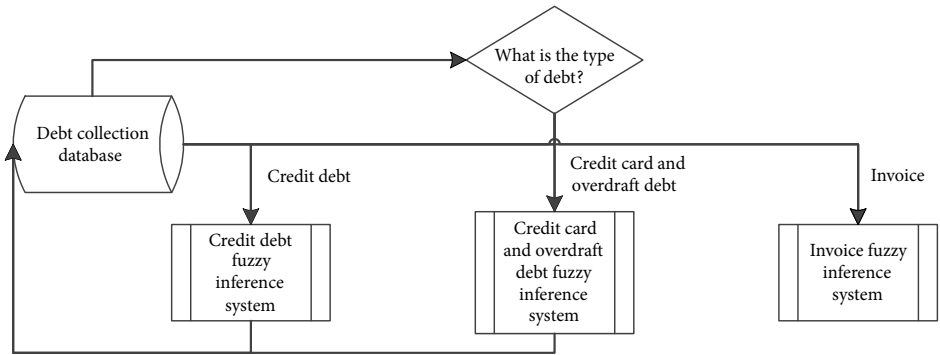
Figure 2. Early debt collection system

inference system. The communication channel is selected accordance to this inference system result. After communicating with the debtor, the response of the debtor is also added to the debt collection database.

3.1. Early collection system design for credit debts

The inputs of fuzzy inference systems for credit debts are *amount of loan*, *wealth of debtor*, *amount of other debts*, *past history of debtor*, and *active customer since*. Each input has a membership function composed of linguistic variables. We have defined these membership functions using data driven approaches and expert opinions. Past data have been analyzed for each input and the scales and membership functions in Figure 3 have been constructed together with the experts.

The input *amount of loan* is the amount of debt that the debtor owns to the considered bank. The membership functions for *amount of loan* are defined by seven linguistic terms as shown in Figure 3a (1 Turkish Lira (TL) = US\$ 0.44). For instance the membership function for the linguistic term *Very High (VH)* is defined by Eq. (6):

$$\mu_{VH}(x) = \begin{cases} 0,\ x < 200,000\,TL \\ \dfrac{x - 200,000}{40,000},\ 200,000\,TL \le x < 240,000\,TL\ . \\ 1,\ 240,000\,TL \le x \end{cases} \tag{6}$$

The input *Wealth of Debtor* is the amount of total wealth that the debtor has. This includes all the real estates and capital assets owned by the debtor. The membership functions of this input are as in Figure 3b. The input *Amount of Other Debts* refers to the total debt owed by the debtor to the other institutions except the considered bank. We have used exactly the same membership functions as *Amount of Loan* for *Amount of Other Debts* (see Figure 3c). The input *Past History of Debtor* reflects the previous debt payment patterns of the debtor. The scale for this input is defined based on debtor's trouble index. The membership functions of this input are given in Figure 3d. Debtor's trouble index (DTI) is calculated as in Eq. (7).
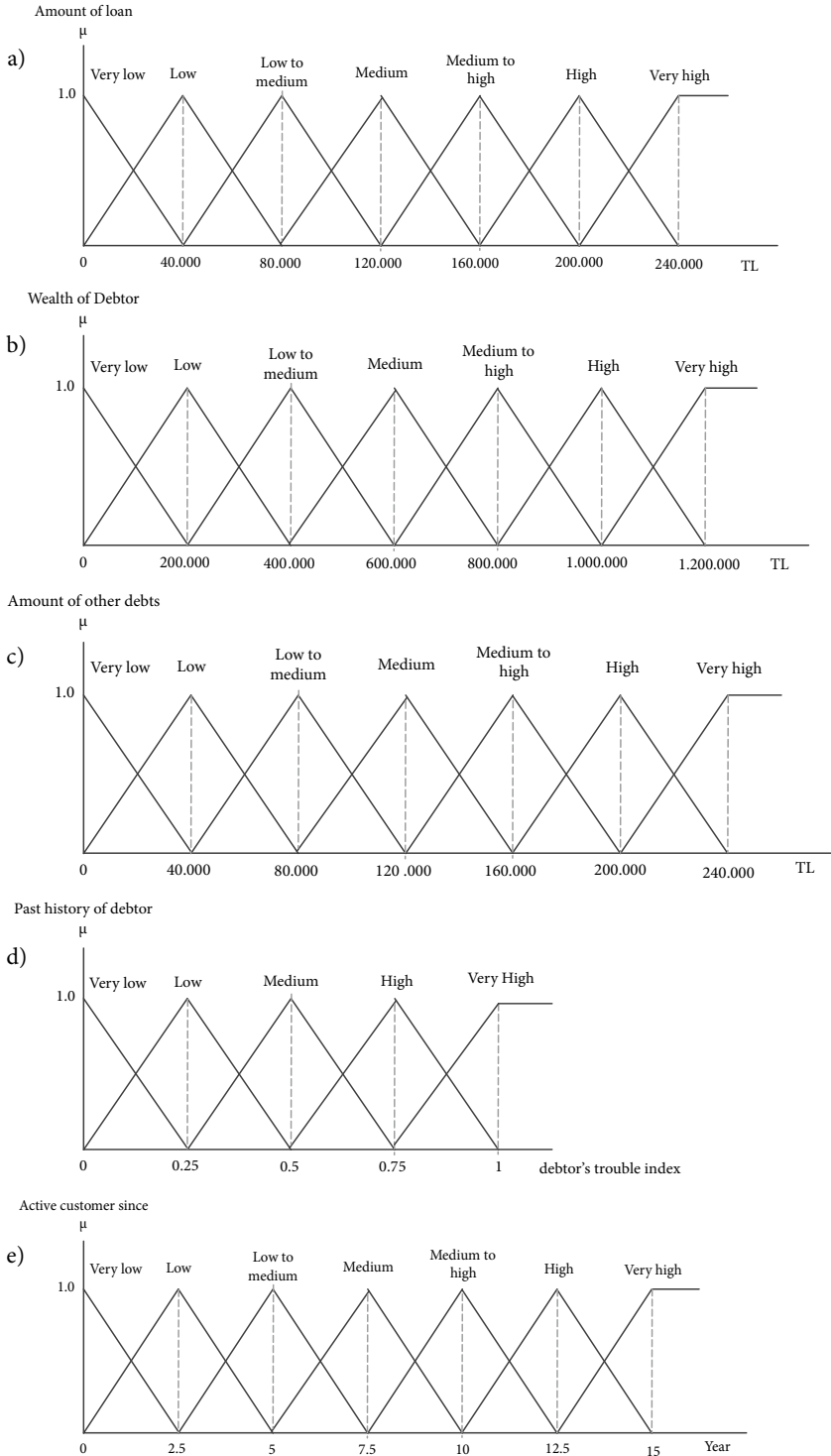
Figure 3. Credit debt inputs and their membership functions

$$DTI = \begin{cases} \dfrac{\sum_{i=1}^{n} Amount\ of\ Delayed\ Payment_i \times Delay\ Time_i}{90 \times \sum_{i=1}^{n} Amount\ of\ Ontime\ Payments_i}, & \sum_{i=1}^{n} Amount\ of\ Ontime\ Payments \neq 0 \\[2em] \dfrac{\sum_{i=1}^{n} Delay\ Time_i}{90}, & \sum_{i=1}^{n} Amount\ of\ Ontime\ Payments = 0, \end{cases} \qquad (7)$$

where *Amount of Delayed Payment$_i$* refers to the periodical amount of the debt delayed by the debtor in the ith period (i = 1, 2, …, n); *Amount of Ontime Payments$_i$* refers to the periodical amount of debt paid by the debtor on time in the ith period (i = 1, 2, …, n); *Delay Time$_i$* refers to the total number of days that the debtor delayed his/her payments in the ith period.

The input *Active customer since* shows the past duration since the debtor is the active customer of the considered bank (see Figure 3e).

The debt collection agency indicated that data records for the inputs *Wealth of debtor* and *Active customer since* may not be available for every debtor. Hence, our proposed FIS system can work even no data for these inputs exist. Considering all the inputs, a total of 15.680 rules have been generated.

A part of these rules obtained by Matlab R2009b are given in Figure 4.

For instance, the first rule in Figure 4 is as follows:

If *amount of loan* is medium and *wealth of debtor* is low to medium and *amount of other debts* is low and *past history of debtor* is high and *active customer since* is very low, then *the possibility of repaying debt* is low.

In Figure 4 only three rule outputs have shaded areas which are the outputs of the rules 4, 13 and 14. The other rule outputs have zero membership degree since red colored input values do not have any intersection with at least one of the input membership functions. Finally, the aggregated output membership function of the rules 4, 13 and 14 is illustrated as the last figure in the outputs column. This shaded area is defuzzified by center of gravity method and this gives the *possibility of repaying the debt* for a certain debt.
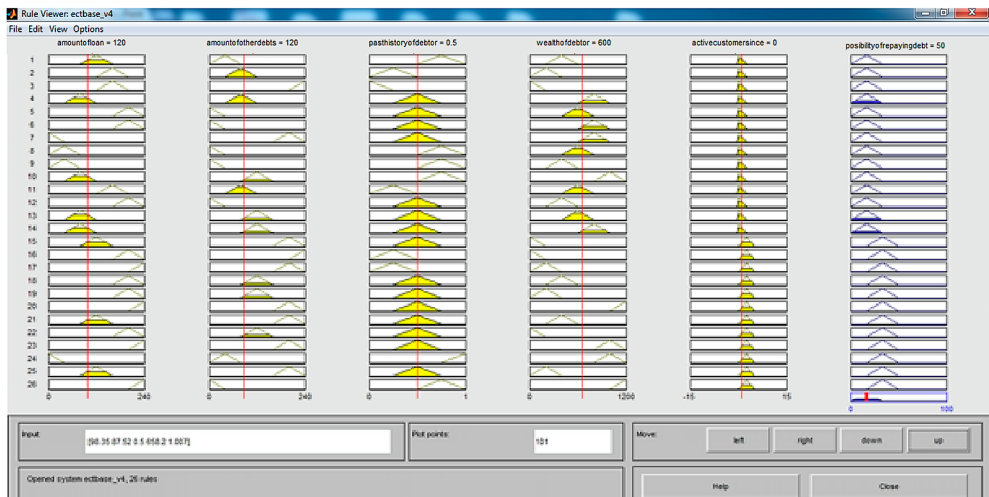

Figure 4. Fuzzy rule sets for credit debts from Matlab interface

The flow chart of the fuzzy rule based inference system for credit debts is given in Figure 5. The five inputs yield *the possibility of repaying debt* and this output determines the subsystem with respect to its level. The flow charts for other debt systems will not be given in this paper since the limited space constraints.

Figure 5 produces seven different subsystems with respect to the possibility levels. We only give the details of subsystems A, D, and E in Figure 6, Figure 7, and Figure 8, respectively. For instance, in Figure 6, since the *possibility of repaying debt* is very low; the subsystem intends to use lower cost communication tools among activity profiles such as e-mails rather than phone calls. The sub-system also uses the past communication data as an input for selecting the proper activity profile. The previous response of the debtor, its type and the low possibility of repaying debt are the inputs for the initial customer communication channel selection. For instance, if we don't have previous responses from the debtor and we know that the possibility of repaying the debt is low, then we e-mail to the debtor on the third past due day. Until we get a response from the debtor, we continue sending e-mails to the debtors on 5th and 10th days. If the debtor does not respond to e-mails, we send SMS on the 15th, 60th and 80th day until we get a response from the debtor. If the debtor responded to e-mail in our previous communication, we again use the same activity profile. If the debtor responded to SMS in our previous communication, then we send an e-mail to the debtor on the 3rd past due day. Until we get a response from the debtor, we continue sending SMSs to the debtors on 10th, 30th, 60th and 80th days. If the debtor responded only to Telephone call in our previous communications, first we send an e-mail on 3rd day. If no response, we send SMS on 10th day and we call the debtor on 40th and 80th days until we get a response. If we receive a positive response to any communication type that the debtor will pay the debt then we send a positive thankful message, otherwise we send a warning message informing that a collection in litigation will start. All the information regarding the results of communications is stored in Debt Collection Database.

Figure 7 presents the actions that should be taken to the medium possibility of repaying debt. In this figure, we are ready to pay a higher cost with respect to the case in Figure 6.
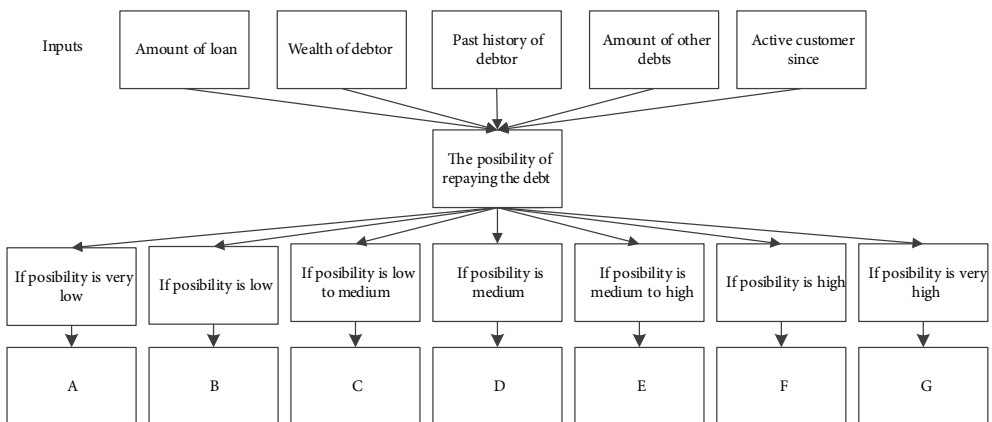

Figure 5. Fuzzy rule based inference system for credit debts

Figure 6. Flow chart Part A- very low possibility

Hence, different activity profiles are defined in Figure 7. For instance, when we compare Figure 6 and Figure 7, one of the differences comes from the communication type on the 20th, 40th, 60th and 80th days. If the debtor responded to SMS in our previous communication, in Figure 7 we send SMSs on 20th, and 40th days and we call the debtor on 60th and 80th days whereas in Figure 6 we send only SMSs on 30th, 60th and 80th past due days.

Figure 8 presents the actions that should be taken to the very high possibility of repaying the debt. In this figure, we are ready to pay a higher cost with respect to the case in Figure 7. Hence, we generally prefer telephone calls for communication.

3.2. Early collection system design for credit card and overdraft debts

The inputs of fuzzy inference systems for credit card and overdraft debts are *amount of loan*, *active customer since, amount of other debts*, *past history of debtor* and *credit limit*. Each input has a membership function composed of linguistic variables. In the same way as the system for credit debts, the past data have been analyzed for each input and the scales and triangular membership values have been constructed together with the experts. Triangular membership values are determined based on both the past data and expert opinions.
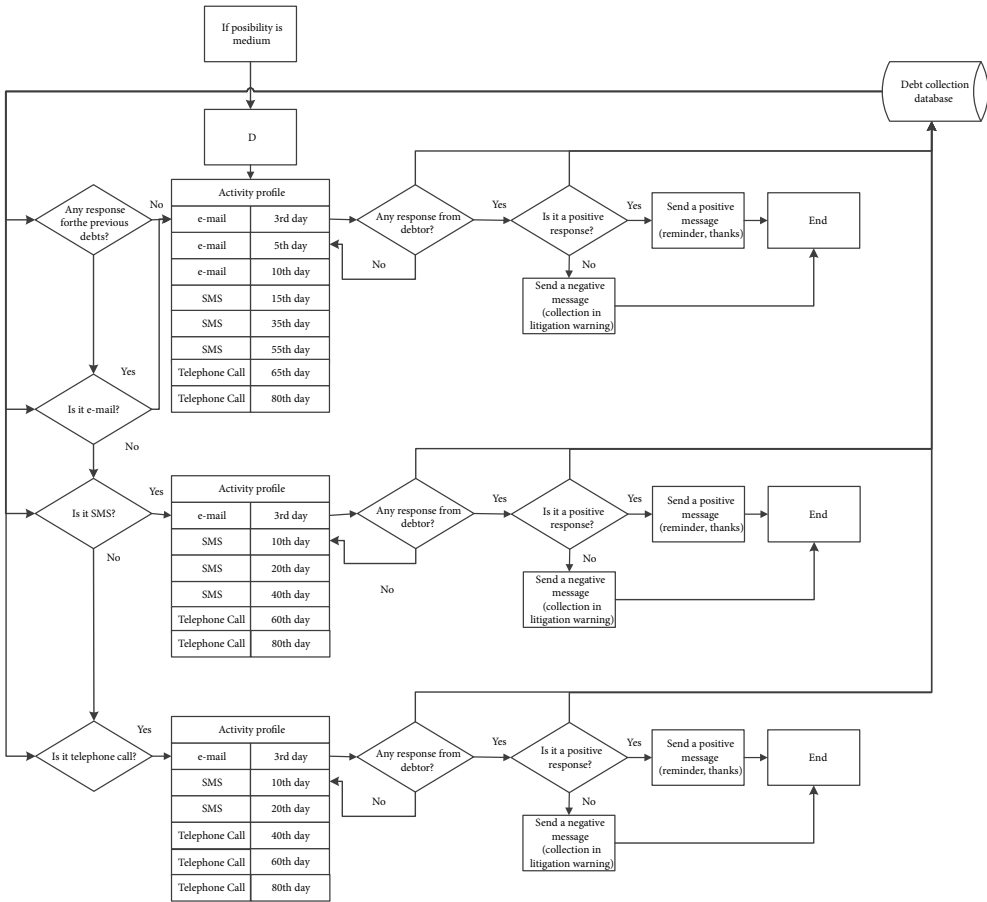
Figure 7. Flow chart Part D- medium possibility

The experts categorized the data by considering the linguistic classes. For instance, the amount of loan for credit cards debts ranges from 0 to 30.000 TL. The experts constructed the category Medium (M) for the debts between 10.000 and 20.000 where the most possible value is 15.000 TL.

The definitions of the inputs *amount of loan, active customer since, amount of other debts,* and *past history of debtor* are the same as it is in the early debt collection system for credit debts. But the triangular membership values for *amount of loan, active customer since,* and *amount of other debts* have been changed as shown in Table 1. The input *credit limit* represents the level of the credit given to a customer with respect to the considered bank's assessment. The input *wealth of debtor* is replaced with the input *credit limit* since credit limit is already defined based on the wealth of debtor.

The debt collection agency indicated that data records for the inputs *amount of other debts* and *active customer since* may not be available for credit card and overdraft debts. The proposed system can work without the inputs of *amount of other debts* and *active customer since*.
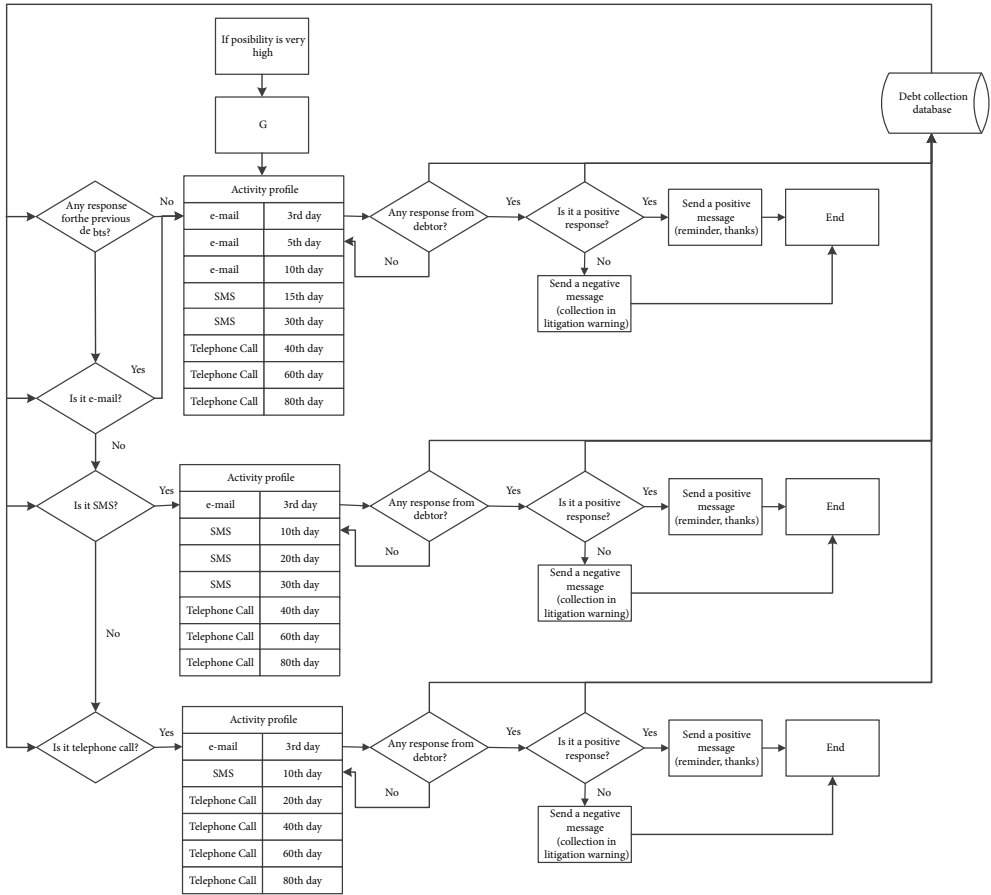
Figure 8. Flow chart Part G- very high possibility

Table 1. Triangular membership values for credit card and invoice debts

	Amount of loan for credit card debts	Amount of loan for invoice debts	Active customer since for both debts	Amount of other debts for credit card debts	Credit limit for credit card debts	Criticality for invoice debts
VL	(0,0,5000)	(0,0,427)	(0,0,1.67)	(0,0,16667)	(0,0,5000)	(0,0,0.167)
L	(0,5000,10000)	(0,427,833)	(0,1.67,3.33)	(0,16667,33333)	(0,5000,10000)	(0,0.167,0.333)
LM	(5000,10000,15000)	(427,833,1250)	(1.67,3.33,5)	(16667,33333,50000)	(5000,10000,15000)	(0.167,0.333,0.5)
M	(10000,15000,20000)	(833,1250,1667)	(3.33,5,6.67)	(33333,50000,66667)	(10000,15000,20000)	(0.333,0.5,0.667)
MH	(15000,20000,25000)	(1250,1667,2083)	(5,6.67,8.33)	(50000,66667,83333)	(15000,20000,25000)	(0.5,0.667,0.833)
H	(20000,25000,30000)	(1667,2083,2500)	(6.67,8.33,10)	(66667,83333,100000)	(20000,25000,30000)	(0.667,0.833,1)
VH	(25000,30000,30000)	(2083,2500,2500)	(8.33,10,10)	(83333,100000,100000)	(25000,30000,30000)	(0.833,1,1)

Considering all the inputs, a total of 15.680 rules have been generated, the same number as in the system for credit debts. As we indicated in Section 4.1, the flow charts for credit card and overdraft debt system will not be given due the space constraints. The basic idea behind this

system is similar to the system for credit debts. Similarly, this system intends to use lower cost communication tools among activity profiles for lower *possibility of repaying debt*. This system also uses the past communication data as an input for selecting the proper activity profile.

3.3. Early collection system design for invoices

Invoice debts are the debts of customers for electricity, cable television, natural gas, mobile phones, etc. The inputs of fuzzy inference systems for invoices are *amount of loan*, *active customer since, past history of debtor* and *criticality*. In the same way as the system for previous systems, the past data have been analyzed for each input and the scales and membership functions have been constructed together with the experts.

The definitions of the inputs *amount of loan, active customer since,* and *past history of debtor* are the same as it is in the credit card and overdraft debts system. But the membership functions for *amount of loan* have been changed as shown in Table 1. *Criticality* refers to the magnitude of the living difficulty in case that service is shut off because of unpaid invoice. For instance, when the electricity is shut off due to the unpaid invoice, the debtor will not be able to realize his/her daily activities such as using refrigerators, internet connection, or heating, which means that the criticality of electricity invoice is very high for the debtors. Since seven symmetrical linguistic terms are used to define the *criticality* input, criticality index (CI) is obtained by dividing a unit to six as shown in Table 1.

The debt collection agency indicated that data records for the input *criticality* may not be available for some of the invoices. Our proposed FIS system for invoices can work without this input. Considering all the inputs, a 1.960 total of rules have been generated. As we indicated in Section 3.1, the flow charts for invoice system will not be given due the space constraints. The basic idea behind this system is similar to the previous debt collection systems. Similarly, this system intends to use lower cost communication tools among activity profiles for lower *possibility of repaying the invoice*.

3.4. Performance evaluation of proposed early debt collection system

The experts of the project are one professor and two assistant professors from ITU Industrial Engineering Department; two managers, two business analysts, and one software developer from the debt collection agency. All the experts have an experience of at least 15 years. Each of the inference system's rules has been discussed carefully by this expert team. After the system is completed the debt collection agency provided a sample data of 250 cases per each debt type for testing the developed system. When we match the output of the fuzzy inference system (expected result) with the sample data belonging to invoice debts, Table 2 has been obtained. In this table, predicted *positive* means that the score for *possibility of paying debt* is higher than 50; and predicted *negative* for lower than 50; actual *positive* means that the debtor paid his/her debt; and actual *negative* means that the debtor did not pay his/her debt.

To evaluate the performance of the system we used accuracy and precision measures which are calculated by Eq. (8) and Eq. (9) (Tan *et al.* 2006):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}; \tag{8}$$

Table 2. Performance evaluation of invoice debts system

		Predicted		
		Positive	Negative	Total
Actual	Positive	118 (TP)	15 (FN)	133
	Negative	21 (FP)	96 (TN)	117
	Total	139	111	250

$$\text{Precision} = \frac{TP}{TP + FP}, \qquad (9)$$

where *TP*, *TN*, *FP* and *FN* represent true positive, true negative, false positive and false negative, respectively. From Table 2, the following performance measurements have been obtained:

$$\text{Accuracy} = \frac{214}{250} = 85.64\%;$$

$$\text{Precision} = \frac{118}{139} = 84.9\%.$$

For the other two inference systems the following results have been obtained: For credit debts FIS, *Accuracy* is 80.4% and *Precision* is 83.6%. For credit card and overdraft debts FIS, *Accuracy* is 82.7% and *Precision* is 84.6%. The accuracy and precision values proved that the prediction of the fuzzy inference systems is very high which means that the proposed system works properly.

The outcomes of the system have been very consistent with experts' and managers' expectations. Besides, the debt collection agency using the developed FIS systems informed that this new system significantly reduced debt collection costs. The customer satisfaction level is also increased due to the standard and appropriate communication channel selection.

Three dimensional plots have also proved that the proposed system produced consistent and reliable outcomes. For instance, Figure 9a shows the effects of *amount of loan* and *criticality* on the *possibility of repaying debt* whereas Figure 9b illustrates the effects of *amount of loan* and *active customer since* on the *possibility of repaying debt*. In Figure 9a, very low criticality and very high amount of loan give a very low *possibility of repaying debt*, which can be seen in dark blue colored area. Very high *criticality* and very low *amount of low* yield a very high *possibility of repaying debt* in yellow colored area. In Figure 9b, a very high *amount of loan* and very high *active customer since* (a customer for many years) yield a medium level of *possibility of repaying debt*. In these three dimensional plots the same colored areas represent equal levels of the output. Hence in each of Figure 9a–9b the same color means the equal possibility of repaying debt. Similarly, the three dimensional plots which can be generated for the other two sub-systems have been obtained. The results indicated that all the relations between inputs and output were consistent and reliable.

The developed FIS system handled the complex debt collection problem by dividing the large problem into sub-problems based on the classification of debts. Thus, the agency employees could be sure about which action they should choose when a file arrives at them.

The vagueness about whether an input in a certain class of debts is low, medium, high, etc. has been left to the FIS, saving the employees from the hard decisions. Entering crisp inputs into the developed system simply gives an advice to the employee about his/her action.
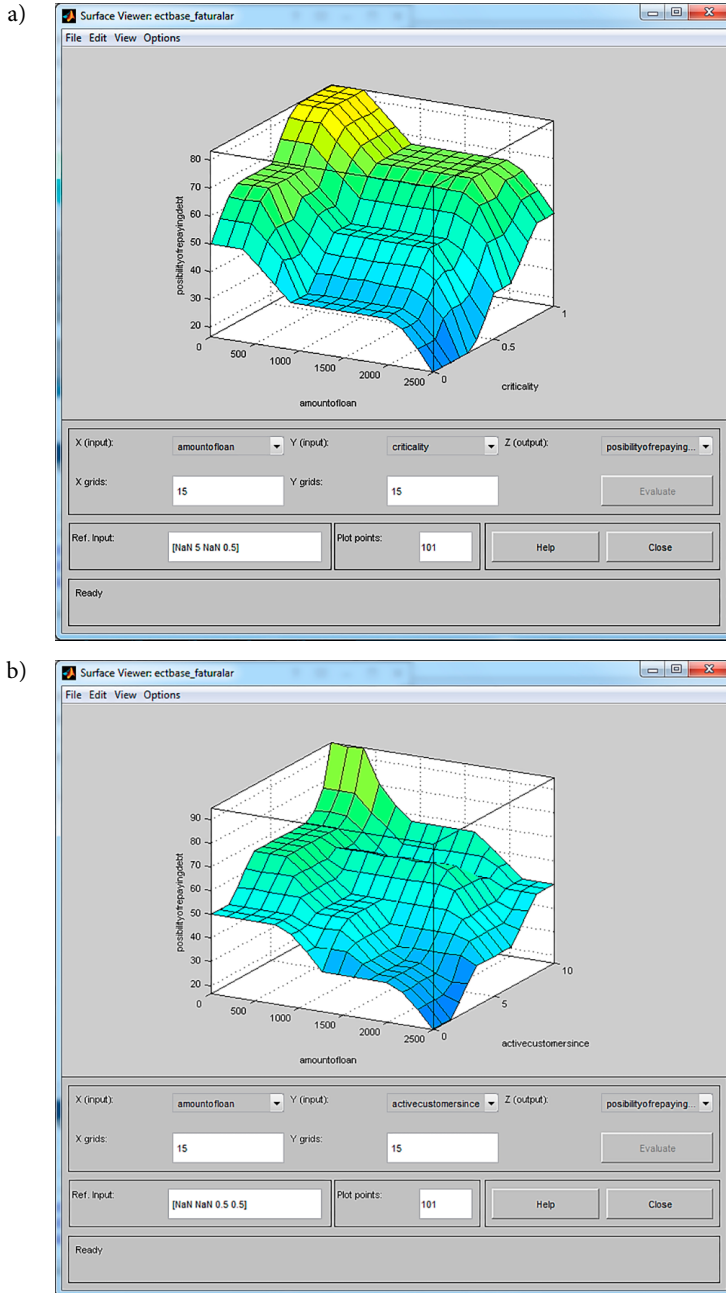

Figure 9. Three dimensional rules for invoice debts

Conclusions

Debt collection systems are vague and complex systems since they try to predict human behaviors related to the repayment of debts. Therefore, fuzzy rule based inference systems are excellent tools to model such problems. We developed an efficient early debt collection system model based on Mamdani-type fuzzy inferences. Mamdani-type fuzzy inference is widely accepted for capturing expert knowledge since it allows us to describe the expertise in more intuitive and more human-like manner. With this model, the communication way with debtors and the possibility of repaying a debt can be easily determined. Considering the types of debts, three different Mamdani-type fuzzy inference sub-systems for debt collection have been developed. This required some different inputs to be used in each system. The main aim in developing each sub-system is to obtain minimum communication cost by taking the past communication experiences of debtors. The flow chart for each-sub system let you follow the correct communication way with the debtors. Three dimensional plots show the effects of any two inputs on the possibility of repaying the debt. The total number of fuzzy rules in the three sub-systems is 33,320 where all the possible combinations are considered. The accuracy and precision measurements of the three FIS systems showed that their performances are very high.

For further research the model can be enhanced by ANFIS (Adaptive Neuro Fuzzy Inference Systems) to generate new rules by improving the existing rules. Since ANFIS integrates both neural networks and fuzzy logic principles and has a learning capability to approximate nonlinear functions, it has potential to capture these benefits in a single framework (Shoorehdeli *et al.* 2009; Sangaiah *et al.* 2015). A similar investigation for early debt collection problem can be handled by fuzzy cognitive maps. In the literature, FIS have been also integrated to fuzzy cognitive maps to define causal relationships. This integrated system is called fuzzy inference cognitive maps (Jones *et al.* 2004). Fuzzy inference cognitive maps can be integrated to early debt collection system to analyze the causal relations.

Acknowledgements

References

Abe, N.; Thomas, V. P.; Kowalczyk, M.; Melville, P.; Pendus, C.; Reddy, C. K.; Jensen, D. L.; Bennett, J. J.; Anderson, G. F.; Cooley, B. R.; Domick, M.; Gardinier, T. 2010. Optimizing debt collections using constrained reinforcement learning, in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, July 25–28, Washington, DC, USA, 75–84. https://doi.org/10.1145/1835804.1835817

Behret, H.; Oztaysi, B.; Kahraman, C. 2011. A fuzzy inference system for supply chain risk management, in Y. Wang, T. Li (Eds.). *Practical applications of intelligent systems*. Berlin, Heidelberg: Springer, 429–438.

Boyacioglu, M. A.; Avci, D. 2010. An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: the case of the Istanbul Stock Exchange, *Expert Systems with Applications* 37(12): 7908–7912. https://doi.org/10.1016/j.eswa.2010.04.045

Cevik Onar, S.; Ates, N. Y. 2008. A fuzzy model for operational supply chain optimization problems, *Multiple-Valued Logic and Soft Computing* 14(3–5): 355–370.

Cevik Onar, S.; Oztaysi, B.; Otay, İ.; Kahraman, C. 2015. Multi-expert wind energy technology selection using interval-valued intuitionistic fuzzy sets, *Energy* 90(1): 274–285. https://doi.org/10.1016/j.energy.2015.06.086

Chen, S. C.; Huang, M. Y. 2011. Constructing credit auditing and control & management model with data mining technique, *Expert Systems with Applications* 38(5): 5359–5365. https://doi.org/10.1016/j.eswa.2010.10.020

Chin, A. G.; Kotak, H. 2006. Improving debt collection processes using rule-based decision engines: a case study of Capital One, *International Journal of Information Management* 26(1): 81–88. https://doi.org/10.1016/j.ijinfomgt.2005.10.002

FDSPA. 2014. *Fair Debt Collection Practices Act* [online], [cited 20 May 2015] Available from Internet: http://files.consumerfinance.gov/f/201403_cfpb_fair-debt-collection-practices-act.pdf

Fei, C. I. 2010. Evaluate the performance of cardholders' repayment behaviors using artificial neural networks and data envelopment analysis, in *Sixth International Conference on Networked Computing and Advanced Information Management (NCM)*, 16–18 August 2010, Seoul, Korea, 478–483.

Georgopoulos, E. F.; Giannaropoulos, S. M. 2007. Solving resource management optimization problems in contact centers with artificial neural networks, in *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, 29–31 October 2007, Patras, Greece, 405–412.

Giovanis, E. 2012. Study of discrete choice models and adaptive neuro-fuzzy inference system in the prediction of economic crisis periods in USA, *Economic Analysis and Policy* 42(1): 79–95. https://doi.org/10.1016/S0313-5926(12)50006-8

Hector, C. 2011. Debt collection in the information age: new technologies and the fair debt collection practices act, *California Law Review* 99(6): 1601.

Howard, E. 2012. Survey forecasts debt collection crisis: big issues and increase in debt cases – and no clear remedy, *UK Business Today* 33(7/8): 667.

Huls, N. 1992. American influences on European consumer bankruptcy law, *Journal of Consumer Policy* 15: 125–142. https://doi.org/10.1007/BF01352132

IBISworld. 2014. *Debt collection agencies market research report* [online], [cited 16 May 2015]. Available from Internet: http://www.ibisworld.com/industry/default.aspx?indid=1474

Jee, T. L.; Tay, K. M.; Lim, C. P. 2015. A new two-stage fuzzy inference system-based approach to prioritize failures in failure mode and effect analysis, *IEEE Transaction on Reliability* 64(3): 869–877. https://doi.org/10.1109/TR.2015.2420300

Jones, P. M.; Roy, R.; Corbett, J. 2004. Fuzzy Information, in *Processing of IEEE Annual Meeting of the NAFIPS'04*, 533–538.

Kahraman, C.; Beskese, A.; Kaya, I. 2010. Selection among ERP outsourcing alternatives using a fuzzy multi-criteria decision making methodology, *International Journal of Production Research* 48(2): 547–566. https://doi.org/10.1080/00207540903175095

Kaya, I.; Oztaysi, B.; Kahraman, C. 2012. A two-phased fuzzy multicriteria selection among public transportation investments for policy-making and risk governance, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 20: 31–48. https://doi.org/10.1142/S021848851240003X

Lund, S. 2010. Soft debt collection – collecting money without alienating your customer, in J. Reuvid (Ed.). *The Business guide to credit management: advice and solutions for cost control, financial risk management and capital protection*. Kogan Page, Ltd., 119–125.

Oztaysi, B.; Behret, H.; Kabak, O.; Uçal Sarı, I.; Kahraman, C. 2013. Fuzzy inference systems for disaster response, in B. Vitoriano, J. Montero, D. Ruan (Eds.). *Decision aid models for disaster management and emergencies*. Paris: Atlantis Pres, 75–94.

Rezvan, P.; Azadnia, A. H.; Noordin, M. Y.; Seyedi, S. N. 2014. Sustainability assessment methodology for concrete manufacturing process: a fuzzy inference system approach, *Advanced Materials Research* 845: 814–818.

Ross, T. J. 1995. *Fuzzy logic with engineering applications*. Addison Wesley.

Sangaiah, A. K.; Thangavelu, A. K.; Gao, X. Z.; Anbazhagan, N.; Durai, M. S. 2015. An ANFIS approach for evaluation of team-level service climate in GSD projects using Taguchi-genetic learning algorithm, *Applied Soft Computing* 30: 628–635. https://doi.org/10.1016/j.asoc.2015.02.019

Shoorehdeli, M. A.; Teshnehlab, M.; Sedigh, A. K. 2009. Training ANFIS as an identifier with intelligent hybrid stable learning algorithm based on particle swarm optimization and extended Kalman filter, *Fuzzy Sets and Systems* 160(7): 922–948. https://doi.org/10.1016/j.fss.2008.09.011

Takahashi, M.; Tsuda, K. 2013. Towards early detections of the bad debt customers among the mail order industry, in T. Matsuo, R. C. Palacios (Eds.). *Electronic business and marketing: new trends on its process and applications*, 167–176.

Tan, P. N.; Steinbach, M.; Kumar, V. 2006. *Introduction to data mining*. Boston, USA: Pearson International Edition.

Tavana, M.; Azizi, F.; Azizi, F.; Behzadian, M. 2013. A fuzzy inference system with application to player selection and team formation in multi-player sports, *Sport Management Review* 16(1): 97–110. https://doi.org/10.1016/j.smr.2012.06.002

Tay, K. M.; Lim, C. P. 2008a. On the use of fuzzy inference techniques in assessment models. Part I – Theoretical Properties*, Fuzzy Optimization and Decision Making* 7(3): 269–281.

Tay, K. M.; Lim, C. P. 2008b. On the use of fuzzy inference techniques in assessment models. Part II – Industrial applications, *Fuzzy Optimization and Decision Making* 7(3): 283–302.

Vecchio, M. D.; Jin, S.; Mistretta, A.; Rolando, H.; Tuck, H. 2006. *Designing a search mechanism for debt collection in Systems and Information Engineering Design Symposium IEEE*, April 26–28, Charlottesville, VA, 168–173.

Wang, H. Y.; Liao, C.; Kao, C. H. 2013. A credit assessment mechanism for wireless telecommunication debt collection: an empirical study, *Information Systems and e-Business Management* 11(3): 357–375. https://doi.org/10.1007/s10257-012-0192-x

Yaici, W.; Entchev, E. 2016. Adaptive Neuro-Fuzzy Inference System modelling for performance prediction of solar thermal energy system, *Renewable Energy* 86: 302–315. https://doi.org/10.1016/j.renene.2015.08.028